Proceedings of the Indian National Science Academy (2026) 92:1259–1274
https://doi.org/10.1007/s43538-025-00479-y

RESEARCH PAPER

Rice yield prediction using UAV-mounted RGB sensors and machine learning algorithms

Rahul Tripathi[1] · Ashish Kumar Gouda[1] · Shiv Sundar Jena[1] · Rajshri Rajlaxmi Mohapatra[1] · Milan Kumar Lal[1] · Sushanta Kumar Dash[1] · Rabi Narayan Sahoo[2] · A. K. Nayak[1]

Received: 8 January 2025 / Accepted: 4 June 2025 / Published online: 7 July 2025
© Indian National Science Academy 2025

Abstract
Rice yield prediction before harvest enables farmers in optimizing the risk of yield loss and adoption of suitable agronomic management practices. Remote sensing data captured using low cost RGB sensors mounted on drones offers a reliable alternative to conventional methods. This study was conducted at ICAR-Central Rice Research Institute, Cuttack, Odisha, India using different rice cultivars, with the objective of developing machine learning (ML) models for rice yield prediction using predictor variables (vegetation indices) generated from data captured using RGB sensors. From 25 vegetation indices (VIs) derived during the panicle initiation stage via RGB imagery, 17 were selected for model development using the variance inflation factor (VIF) technique. Using the selected vegetation indices as a predictor variable, four machine learning models gradient boosting machine (GBM), random forest (RF), artificial neural network (ANN), and support vector machine (SVM) were developed with rice yield as the dependent variable. Among all model, the RF model achieved best prediction accuracy in predicting rice yield (RMSE = 0.27 t ha^{-1}). The RF model predicted yield in the range of 3.16–5.50 (t ha^{-1}). The coloration (CI), hue (H) and intensity (I) vegetation indices were identified as the most important contributors in accurate prediction of rice yield across all the ML models. The study demonstrated that the proposed approach is reliable and cost-effective method of the of rice grain yield prediction however, models need multi-season and multilocational validation before suggesting a large-scale application.

Keywords Drone · Rice yield prediction · Sensors · Variance inflation factor · Random forest

✉ Rahul Tripathi
rahulcrri@gmail.com

Ashish Kumar Gouda
ashishgouda600@gmail.com

Shiv Sundar Jena
shivsundarjena89@gmail.com

Rajshri Rajlaxmi Mohapatra
rajshrirajlaxmimohapatra@gmail.com

Milan Kumar Lal
milan.lal@icar.gov.in

Sushanta Kumar Dash
skdash139@gmail.com

Rabi Narayan Sahoo
rnsahoo.iari@gmail.com

A. K. Nayak
aknayak20@gmail.com

[1] ICAR-Central Rice Research Institute, Cuttack, Odisha, India

[2] ICAR- Indian Agricultural Research Institute, New Delhi, India

Introduction

With the growing global population, the pressure is escalating to ensure food security. Rice, which serves as the staple food for more than half of the world's population, plays a crucial role in meeting this demand (Chang 2024; Chu and Yu 2020). In India, rice is crucial, providing sustenance to more than 60% of the population, with the cultivated area increasing from 30.81 million hectares in 1950–51 to 43.86 million hectares in 2020–21 (Das et al. 2023). However, the challenge of feeding a growing population is intensified by climate change, which affects water, land, and labor resources, requiring sustainable agricultural practices to boost production while minimizing environmental impacts (Tanaka et al. 2023).

Rice yield prediction is essential not only at national scale, but also for smallholder farmers, as any reduction in yield significantly impact food security, livelihoods, and incomes (Hertel et al. 2010; Verma et al. 2021). For policy

level decision planning at village level, accurate and timely crop yield forecasts are essential. The yield predictions also support crop insurance schemes in assessing the risks posed by climate variability (Zhou et al. 2017; Murthy et al. 2022). Rice yield forecasting is a critical tool to improve agricultural policies and crop insurance schemes (Zhou et al. 2017; Murthy et al. 2022), yet traditional field surveys are labor-intensive and inefficient in resource utilization (Ali et al. 2022). Recent advances in satellite-based remote sensing and Unmanned Aerial Vehicle (UAV) technologies have enabled cost-effective and high-resolution yield predictions across large agricultural areas (Mariotto et al. 2013; Noureldin et al. 2013; Herzig et al. 2021; Zhou et al. 2023). UAVs, in particular, when integrated with machine learning techniques, can improve yield prediction by capturing detailed vegetation indices (VIs), including the Normalized Difference RedEdge Index (NDRE) and Normalized Difference Vegetation Index (NDVI). This method has proven effective for a variety of crops, including wheat, barley and rice (Gong et al. 2018; Maimaitijiang et al. 2019).

Among the available sensers, RGB (Red, Green, Blue) sensors mounted on UAVs offers a more accessible and affordable alternative to multispectral sensors. These sensors are effective for yield prediction, as they can capture key crop attributes like canopy structure, growth stages, and chlorophyll content (Kefauver et al. 2017) RGB cameras are also cost-effective, widely available, and less affected by adverse weather conditions, making them highly suitable for small-scale operations (Bagheri et al. 2013; Gong et al. 2018). UAVs equipped with RGB sensors are capable of providing a practical and efficient tool for monitoring important crop characteristics such as plant height, chlorophyll content and biomass, offering a valuable solution for smallholder farms (Gong et al. 2018; Maimaitijiang et al. 2019).

In modern precision agriculture, machine learning models such as artificial neural networks (ANN) (Fieuzal et al. 2017; Amaratunga et al. 2020), multivariate linear regression, and statistical models (Geipel et al. 2014; Zhou et al. 2017; Gong et al. 2018; Peng et al. 2019) have become integral to predicting crop yields. Among these, models like support vector machine (Kuwata and Shibasaki 2016; Liu et al. 2022a, b) and random forest (Breiman 2001; Li et al. 2020; He et al. 2021) have shown exceptional promise due to their high predictive accuracy and flexibility in handling non-linear data. These machine learning models along with sensors mounted on drone are now seen as powerful tools for yield estimation in diverse crop types, playing a key role in addressing challenges posed by climate change and limited resources (Jhajharia and Mathur 2023; Shahhosseini et al. 2020; Wang et al. 2021; Liu et al. 2022a, b). With the above background, the specific objectives of this study were to derive vegetation indices from RGB sensors and develop machine learning model using these indices for rice yield prediction.

Material and methods

Site description

This study was undertaken at ICAR-Central Rice Research Institute, Cuttack, Odisha, located in eastern India (20°47′N to 85°87′E) (Fig. 1). The on station experimental site is approximately 50 m above sea level having a sub-humid tropical climate, receiving around 1500 mm of annual rainfall with an average annual temperature of 27.6 °C.

Experiment details

A field experiment was conducted with 27 contrasting rice varieties during Jan–May 2024 (*Rabi season*). The nursery for the experiment was sown on 21st January 2024 and transplanting was carried out on 13th February 2024. Variety to Variety isolation distance was 0.6 m. The varieties with different crop durations (Early to mid-early), varying yield potential and plant height were taken. The experiment was conducted in a Completely Randomized Block Design

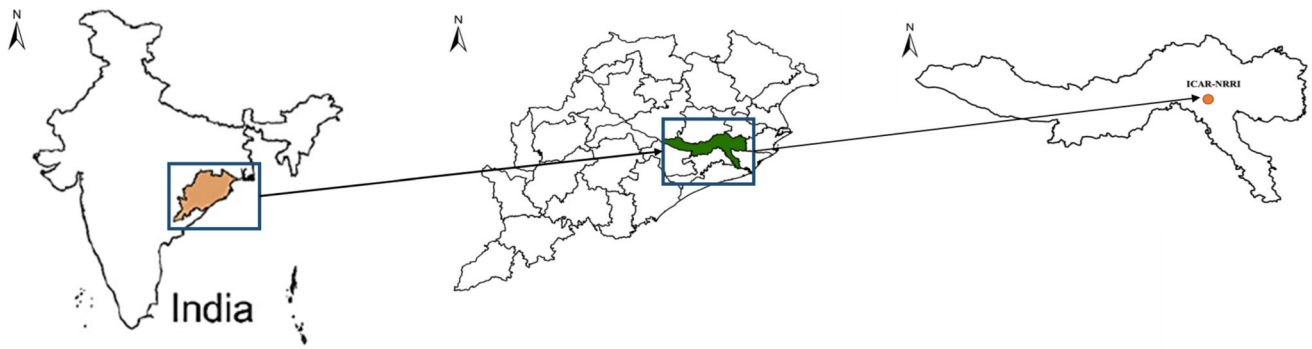


Fig. 1 Location of Study site in eastern India

(CRBD) with three replications and each plot size of around 18 m^2. The recommended fertilizer dose of 100:40:40 kg ha^{-1} (N: P_2O_5: K_2O) was applied and irrigation was provided as per crop requirement during different crop growth stages. Additionally, another experiment was conducted with varying nitrogen doses (0 $kgha^{-1}$, 20 kg ha^{-1}, 40 $kgha^{-1}$, 60 $kgha^{-1}$, 80 kg ha^{-1}, 100 kg ha^{-1}, 120 kg ha^{-1} and 140 kg ha^{-1}) using two rice varieties, Pooja and Swarna (the experimental detail can be accessed from Tripathi et al. 2024).

Image acquisition using unmanned aerial vehicle (UAV)

The IZI, M3 Pro THS UAV, mounted with CMOS sensor was used for capturing 20-megapixel RGB images. Flight duration of UAV was around 25–30 min. The RGB images were acquired synchronising with the panicle initiation stage of each variety (Table 1). The flight plan was prepared using Litchi (Version 4.26.9) software and the images were acquired under clear sky condition, with a view angle of 90^0 from the vertical (NADIR view, at 20 m height above the ground).

Table 1 Date of drone image acquisition for different varieties

Sl. No.	Varieties	Date of image acquisition
1.	CR Dhan 200	10/04/2024
2.	CR Dhan 201	10/04/2024
3.	CR Dhan 202	10/04/2024
4.	CR Dhan 203	10/04/2024
5.	CR Dhan 204	10/04/2024
6.	CR Dhan 205	15/04/2024
7.	CR Dhan 206	10/04/2024
8.	CR Dhan-207	10/04/2024
9.	CR Dhan 209	15/04/2024
10.	CR Dhan 210	15/04/2024
11.	CR Dhan 306	15/04/2024
12.	CR Dhan 307	09/05/2024
13.	CR Dhan 314	09/05/2024
14.	CR Dhan 316	09/05/2024
15.	CR Dhan 602	15/04/2024
16.	Annada	10/04/2024
17.	Naveen	10/04/2024
18.	TN-1	10/04/2024
19.	CR Dhan 324	10/04/2024
20.	CR Dhan 312	10/04/2024
21.	CR Dhan 317	10/04/2024
22.	CR Dhan 805	10/04/2024
23.	CR Dhan 807	10/04/2024
24.	Vandana	10/04/2024
25.	Sahabhagi dhan	10/04/2024
26.	Tapaswini	10/04/2024
27.	CR Dhan 326	09/05/2024

Selection of variables for model building

A literature review was conducted to identify VIs suitable for predicting the crop yield and other crop biophysical parameters. A total of 25 vegetation indices were derived using the three bands of RGB image. VIF technique was followed for identifying the 17 most relevant vegetation indices suitable for predicting the yield. Other studies using these 17 indices for predicting yield and other biophysical parameters are also listed in Table 2. The derived equations for the vegetation indices from RGB images are mentioned in Table 3. The number of input variables plays a crucial role in determining the effectiveness of ML models.

Minimizing the number of input variables improve the reliability and usability of ML models by reducing their overfitting. (Kuhn and Johnson 2013; Steyerberg and Vergouwe 2014). In this study, the vegetative indices were chosen using the VIF approach. The VIF-based variable selection method selects the independent variables with the least amount of collinearity between them after assessing the collinearity between the variables (Valdivieso-Ros et al. 2023).

To identify multicollinearity among the predictor variables, VIF was set to a threshold value below 10. Variables that exceeded this threshold were eliminated until every variable that remained had acceptable VIF scores. The final selection of 17 indices out of 25 was based on their VIF values (Kumar et al. 2023).

In this study, variable selection was performed through stepwise VIF analysis using the R pedometric package. Predictor variables that had high R^2 and VIF values were iteratively eliminated (Akinwande et al. 2015; Akbarian et al. 2020; Aranguren et al. 2020; Han et al. 2020). Each iteration reassessed the VIF values of the remaining variables to ensure compliance with the set threshold (set at 10) (Kim 2019). The process continued until all VIFs were below the threshold, and plot-specific VIs were extracted.

Statistical analysis

Descriptive statistics was calculated for all the generated vegetation indices, used in model training. The measures of central tendency and dispersion were analyzed for the selected variables using R software (R version 4.3.2R Core Team 2023).

Machine learning models for prediction of rice yield

In this study, four machine learning models were used. These models included artificial neural networks (ANN), random forests (RF), support vector machines (SVM) and gradient boosting machines (GBM). The rice yield was considered as

Table 2 The indices selected by variable Inflation factor and the plant parameters affecting these indices

Indices	Parameters	Reference
ExR	Grain yield, Chlorophyll estimation, Leaf area index	Das et al. (2016), Zhou et al. (2017), Guo et al. (2020), Zeng et al. (2021)
ExG	Leaf chlorophyll, Grain yield	Wan et al. (2019), García-Martínez et al. (2020), Sumesh et al. (2021), Liu et al. (2024)
GLA	Leaf area index, Yield	Hasan et al. (2019), Naik et al. (2023), Zhang et al. (2024)
GMR	Grain yield, Above ground biomass, Chlorophyll content, Leaf area index	Wang et al. (2013),Özreçberoğlu et al. (2020), Ge et al. (2021a, b), Liu et al. (2021)
SCI	Chlorophyll content	Kawashima and Nakatani (1998), Fernando Sánchez-Sastre et al. (2020)
MGRVI	Leaf Area Index, Above ground biomass, Dry biomass	Saberioon et al. (2014), Bendig et al. (2015), Li et al. (2019), Cen et al. (2019)
I	Total chlorophyll content	Hu et al. (2010), Chen et al. (2020)
GRVI	Leaf area index, Leaf dry biomass, Leaf total nitrogen, Yield	Bendig et al. (2015), Qiu et al. (2022), Sanches et al. (2018), Zeng et al. (2021)
VARI	Yield, Leaf area index, Above ground biomass, Dry biomass, Chlorophyll content	Hunt et al. (2013), Zhou et al. (2017), Li et al. (2018), Cen et al. (2019), Li et al. (2019), Lussem et al. (2018), Widjaja Putra et al. (2018), Pazhanivelan et al. (2023)
NDGBI	Total Chlorophyll content	Kawashima and Nakatani (1998), Liu et al. (2021)
NDIg	Leaf area index, Dry mass, Leaf area index, Chlorophyll and carotenoid pigment	Lee et al. (2013), Widjaja Putra and Soni (2018), Qiao et al. (2022)
RGBVI	Above ground biomass, Dry and fresh biomass, Plant nitrogen concentration, Leaf nitrogen concentration	Bendig et al. (2015), Tilly et al. (2015), Wang et al. (2022), Liu et al. (2023)
b	Above ground biomass, Leaf area index, Total chlorophyll content, Yield	Liu et al. (2019), Maimaitijiang et al. (2019), Yue et al. (2019), Blackmer et al. (1996), Gupta et al. (2013), Wei et al. (2015), Feng et al. (2022), Nakasagga et al. (2022)
g	Dry biomass, Leaf area index, Yield	Gerard and Buerkert (2001), Jia et al. (2007), Lee and lee (2013), Wang et al. (2013), Zeng et al. (2021)
H	Total chlorophyll content, Leaf nitrogen concentration	Zhang et al. (2005)

Table 2 (continued)

Indices	Parameters	Reference
CI	Chlorophyll content	Abdullah et al. (2019)
ExB	Aboveground biomass Plant height, Leaf area index, Leaf nitrogen concentration Plant nitrogen concentration,	Maimaitijiang et al. (2019), Qiu et al. (2022) Liu et al. (2023)

ExR Excess red index, *ExG* Normalized excess green index, *GLA* Green Leaf Algorithm, *GMR* Green minus red index, *SCI* Soil colour index, *MGRVI* Modified green red vegetation index, *I* Intensity, *GRVI* Green–Red Vegetation Index, *VARI* Visible Atmospherically Resistant Index, *NDGBI* Normalized Difference Green–Blue Index, *NDIg* Normalized red green difference index, *RGBVI* Red Green Blue Vegetation Index, *B* Normalized Blueness Intensity, *g* Green Percentage Index, *H* Hue, *CI* Coloration indices, *ExB* Excess blue vegetation indices

Table 3 Equation for deriving the vegetation indices

Indices	Formula	Reference
ExR	1.4*r-g	Meyer et al. (1999)
ExG	2*g-r-b	Woebbecke et al. (1995)
GLA	(2 g-r—b)/(2 g+r+b)	Louhaichi et al. (2001)
GMR	g-r	Wang et al. (2013)
SCI	(R-G)/(R+G)	Shigeto and Nakatani (1998)
MGRVI	$(g^2—r^2)/(g^2+r^2)$	Bendig, et al. (2015)
I	R+G+B	Zakaluk and Ranjan (2008)
GRVI	(g-r)/(g+r)	Tucker (1979)
VARI	(g-r)/(g+r-b)	Gitelson et al. (2008)
NDGBI	(G−B)/(G+B)	Shigeto and Nakatani (1998)
NDIg	(R-G)/(R+G+0.01)	Wu (2016)
RGBVI	$(g^2$-b*r)/$(g^2$+b*r)	Bendig et al. (2015)
b	B/(R+G+B)	Niu et al. (2018)
g	G/(R+G+B)	Richardson et al. (2007)
H	atan(2*(R−G−R)/30.5 *(G−B))	Escadafal et al. (1994)
CI	(r-b)/r	Escadafal et al. (1994)
ExB	1.4b-g	Qi et al. (1994)

ExR Excess red index, *ExG* Normalized excess green index, *GLA* Green Leaf Algorithm, *GMR* Green minus red index, *SCI* Soil colour index, *MGRVI* Modified green red vegetation index, *I* Intensity, *GRVI* Green–Red Vegetation Index, *VARI* Visible Atmospherically Resistant Index, *NDGBI* Normalized Difference Green–Blue Index, *NDIg* Normalized red green difference index, *RGBVI* Red Green Blue Vegetation Index, *B* Normalized Blueness Intensity, *g* Green Percentage Index, *H* Hue, *CI* Coloration indices, *ExB* Excess blue vegetation indices, *r* Normalized red band, *G* Green band, *B* Blue band, *R* Red band

the dependent variable, with the vegetation indices serving as the predictor variables. R software was used for model building and its assessment.

Random forest (RF) model

The Random Forest (RF) algorithm is an ensemble-based machine learning technique that constructs multiple decision trees during the training phase and aggregates their outputs to enhance predictive accuracy and robustness (Breiman 2001). In addition to creating and aggregating several decision trees, the Random Forest (RF) technique is useful for calculating underlying correlations between the predictor and target variables. In this study, the randomForest package in R was used to build the RF model by adjusting its hyperparameters including splitting at each node (mtry) and the number of trees (ntree) were two important hyperparameters that were modified to improve the model's performance. Approximately 80% of the dataset was allocated for training the model, and the remaining 20% was used for validation. The mtry value was optimized using the randomForest package having tuneRF function, which minimized the Out-of-Bag (OOB) error to provide a reliable in-built estimate of the model's performance. The parameters for the optimally tuned RF model were set to optimal number of trees (ntree), as it resulted in lower error rates after hypertuning of the model.

Gradient boosting machine (GBM) model

Gradient Boosting Machines (GBM) is an ensemble learning technique that builds predictive models sequentially by minimizing errors from prior models. Each tree is built sequentially to predict or adjust for the residual errors of preceding model ensuring the new tree improves upon the previous one (Charoen-Ung & Mittrapiyanuruk 2018). GBM trains the newly added weak classifier as per the negative gradient of the loss function of the current model (Dev and Eden 2019). Moreover, the algorithm's versatility supports both regression and classification tasks, which makes a valuable tool in diverse decision-making contexts. The GBM model was trained using the "caret" package and tuned using "mlr" package and "tune ()" function.

Artificial neural network (ANN) model

Artificial Neural Network (ANN) technique is based on the structure and functioning of the biological nervous system (Gandhi et al. 2016a, b). The term "network" in ANN' refers to the inter–connections between the neurons in the different layers of each system that are able to modify their internal structure in relation to a function objective (Dahikar and

Rode 2014). Their ability to reconstitute the fuzzy rules governing the ideal solution for nonlinear problems makes them particularly suitable for solving such types of challenges. In this study, the ANN model developed using the nnet package was used to create the ANN model. Several hyperparameters like hidden layer's size (number of neurons) and weight decay (decay) were adjusted to build the ANN model. To evaluate the performance of each hyperparameters configuration, cross-validation was carried out using a grid of potential values. The cross-validated performance was used to determine the optimal set of hyperparameters.

Support vector machine (SVM) model

The support vector machine is a robust and flexible ML technique capable of handling both regression and classification tasks. The "e1071" package in R was used to develop the SVM model, utilizing the radial basis function (RBF) kernel for the effectiveness in capturing intricate data patterns. Hyperparameter optimization for cost and gamma was performed through a grid search approach to ensure optimal model performance.

Model assessment and evaluation

The performance of four machine learning models was assessed using four metrics i.e., mean absolute percent error (MAPE), mean absolute error (MAE), coefficient of determination (R^2), root mean square error (RMSE) and normalized root mean-square error (NRMSE) were used

$$R^2 = \frac{\left(\Sigma\left(y_i - \overline{y}\right)^2 - \Sigma\left(\hat{y}_i - y_i\right)^2\right)}{\Sigma\left(y_i - \overline{y}\right)^2} \quad (1)$$

$$\mathrm{RMSE} = \sqrt{\frac{\Sigma\left(\hat{y}_i - y_i\right)^2}{n}} \quad (2)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|yi - \widehat{yi}\right| \quad (3)$$

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{yi - \widehat{yi}}{yi}\right| \quad (4)$$

$$\mathrm{NRMSE} = \frac{RMSE}{\overline{y}} \times 100\% \quad (5)$$

Here, y_i=Actual observed yield of rice, $\hat{y}_i$=Predicted yield of rice by the model, $\bar{y}$=Mean value of the rice yield (observed), n=Total number of samples points and Σ=Summation over the specified values.

Estimation of the variable importance

Vegetation indices are essential in machine learning models, offering vital insights into vegetation health, including biomass and chlorophyll levels, which support accurate ecological forecasting. By capturing critical spatial and temporal patterns, these indices improve model performance in assessing risks (Ferchichi et al. 2022). Their integration enables consistent vegetation monitoring, aiding data-driven decisions in environmental management. Variable importance was calculated across all the models by evaluating the Gini impurity index, root mean squared error and the impact of each variable on the predictive performance. The variables with lower Gini impurity values were given more importance while splitting a node across decision trees.

Generating the yield prediction map

The yield maps for the test site, which were not involved in model development were created using the RF, GBM, SVM, and ANN models through the predict function in the Raster package in R. The methodology is outlined in the flowchart (Fig. 2).

Result and discussion

Descriptive statistics

A measure of variability for each of the predictor variables used in this study is provided in Table 4. Substantial variation was observed in the vegetation indices as well as grain yield across the different rice varieties. The difference between the maximum (6.80 t ha^{-1}) and minimum (2.40 t ha^{-1}) rice yield across the studied varieties indicates the high variability. In the case of VI, the highest CV was observed for ExR (1259.96%), followed by GMR (92.04%), indicating considerable variation across the treatments, while RGBVI recorded the lowest CV i.e., 6.09%. Among the 17 indices, 7 indices (ExR, GMR, SCI, MGRVI, GRVI, VARI and NDIg) were highly variable whereas 6 indices (ExG, GLA, I, NDGBI, CI and ExB) were moderately variable and three indices (RGBVI, b and H) recorded lower variability as per the classification suggested by Wilding (1985).

Vegetation indices (VI) are valuable tools for predicting crop yield, as they are highly sensitive to crop growth and environmental stress. VIs enhances remote sensing image interpretation through linear or nonlinear combinations of spectral bands, playing a crucial role in crop growth monitoring by capturing variations in plant traits like chlorophyll content, canopy structure, and LAI (Li et al. 2014; Xue and Su 2017; Volpato et al. 2021; Tian et al. 2022). Several

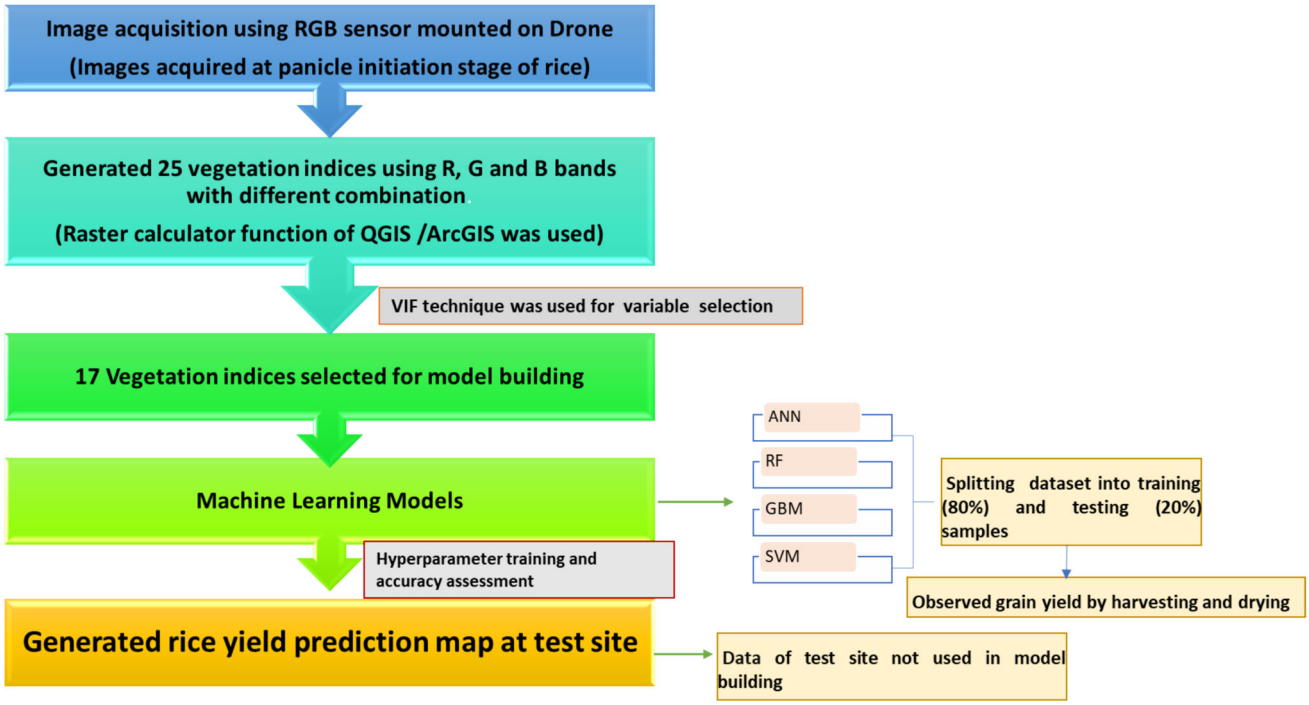


Fig. 2 Flow diagram explaining the overall methodology for yield prediction using drone mounted RGB sensor. RF (Random Forest), GBM (Gradient Boosting Machine), SVM (Support Vector Machine), ANN (Artificial Neural Network)

Table 4 Descriptive statistics of the different variables used in the study

Indices/yield	mean	min	max	Standard deviation	CV (%)
ExR	− 0.004	− 0.13	0.17	0.06	1259.96
ExG	0.40	0.08	0.62	0.09	22.88
GLA	0.26	0.06	0.40	0.05	20.72
GMR	0.30	0.04	1.04	0.28	92.04
SCI	− 0.16	− 0.29	0.02	0.06	36.14
MGRVI	0.31	− 0.03	0.51	0.10	34.00
I	292.50	159.25	420.67	59.47	20.33
GRVI	0.16	− 0.02	0.29	0.06	36.14
VARI	0.22	− 0.02	0.36	0.08	35.77
NDGBI	0.20	− 0.02	0.34	0.07	34.00
NDIg	− 0.16	− 0.29	0.02	0.06	36.13
RGBVI	0.79	0.56	0.89	0.05	6.09
b	0.21	0.16	0.28	0.02	9.50
g	0.47	0.36	0.54	0.03	6.48
H	0.68	0.48	0.86	0.07	10.06
CI	0.37	0.12	0.54	0.07	19.91
ExB	− 0.18	− 0.31	0.03	0.05	29.49
Yield (t ha^{-1})	4.64	2.40	6.80	0.94	20.24

ExR Excess red index, *ExG* Normalized excess green index, *GLA* Green Leaf Algorithm, *GMR* Green minus red index, *SCI* Soil colour index, *MGRVI* Modified green red vegetation index, *I* Intensity, *GRVI* Green–Red Vegetation Index, *VARI* Visible Atmospherically Resistant Index, *NDGBI* Normalized Difference Green–Blue Index, *NDIg* Normalized red green difference index, *RGBVI* Red Green Blue Vegetation Index, *B* Normalized Blueness Intensity, *g* Green Percentage Index, *H* Hue, *CI* Coloration indices, *ExB* Excess blue vegetation indices

researchers (Zhang et al. 2020a, b; Fu et al. 2021; Liu et al. 2023) have used VIs for monitoring crop health and yield.

Bai et al. (2022) used two indices ExG and ExR involving visible bands for crop monitoring. The significant variation in vegetation indices (VIs) like ExR, ExG, GMR, SCI, and others shown in Table 3 can be attributed to differences in crop canopy structure, chlorophyll and nitrogen content in rice leaves. A study conducted by Meena et al. (2020) found that indices such as ExR and ExG, which measure excess red and green were highly sensitive to chlorophyll levels and plant vigor, resulting in greater variability. Additionally, indices like MGRVI and VARI, which rely on green and red bands, exhibit variability due to their sensitivity to leaf area and canopy density (Ruwanpathirana et al. 2024).

Building machine learning models

In this study, RF, ANN, GBM and SVM models were developed to predict the rice yield from 17 vegetation indices (predictor variables) (Sharma et al. 2023) (Fig. 3). The non-linear relationship between the target variable (yield) and predictor variables can potentially result in multiple yield predictions for a single vegetation index value and consequently pose challenges in precisely determining yield (Probst et al. 2019; Yang and Shami 2020). Therefore, fine-tuning of hyperparameters was performed in all the Machine learning models (ML), which is required to avoid overfitting, prediction accuracy and ensure systematic

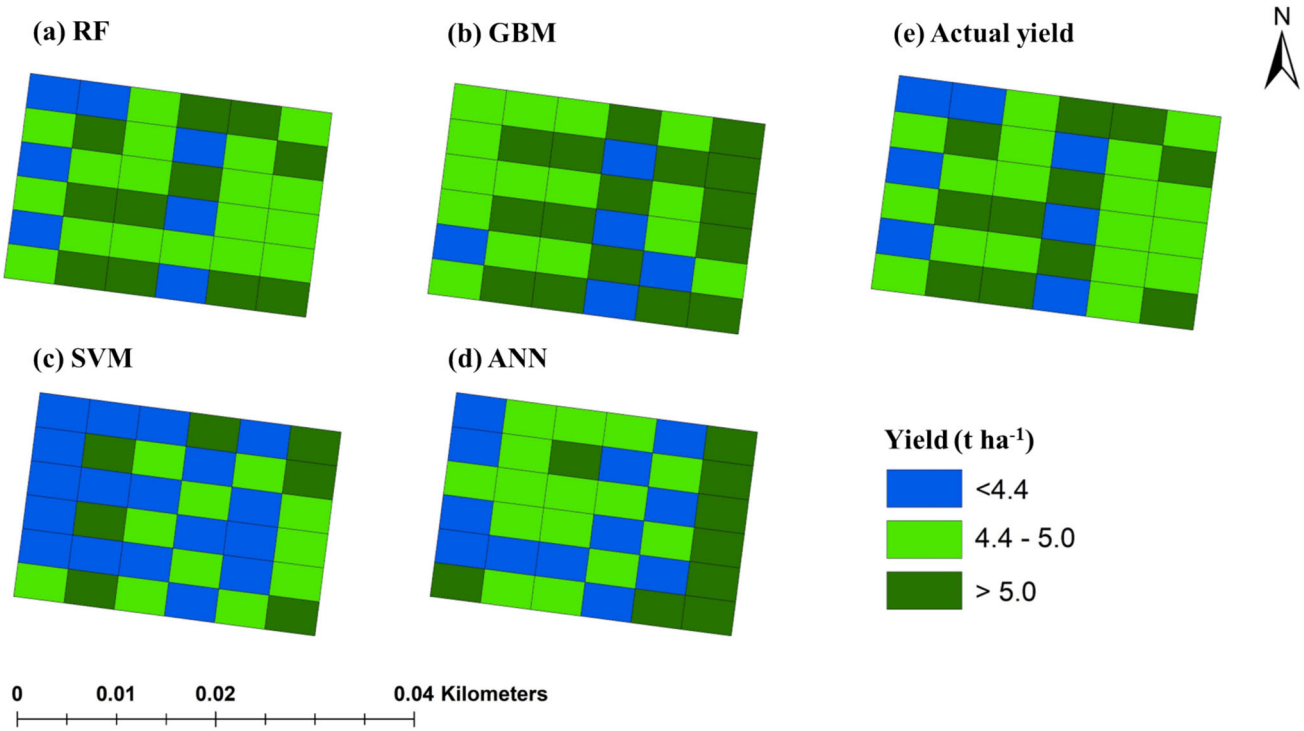


Fig. 3 Prediction maps for the testing plot using **a** RF (Random Forest), **b** GBM (Gradient Boosting Machine), **c** SVM (Support Vector Machine), **d** ANN (Artificial Neural Network) model, **e** Actual yield

representation of the data. The 'randomForest' package in R software was used to develop the RF model (Breiman 2001; Prasad et al. 2021). Bootstrap aggregation (bagging) function (Shahhosseini et al. 2020) was employed to incorporate feature randomness at each node. The model was tuned using 'tuneRF' function with n_{tree} values ranging from 10 to 100 in increments of 10, to determine the most effective model. The best model was identified based on minimum Out-of-Bag (OOB) error (Sitienei et al. 2023) of 0.85, optimal number of n trees as 100 and 'mtry' value was set at 3, considering 3 variables at each split (Breiman 1996; Mohite et al. 2023).

The ANN model was built using "nnet" package (Ripley 2013; Tripathi et al 2024) and was optimized using the caret package. The final model was configured with setting the number of neurons to 3, weight decay values of 0.01 (size = 5), the value of weight decay parameter offset at 0.001 and maxit = 4000 (Bishop 2006; Huber et al. 2024.), a fivefold cross-validation technique was used during model training. (Tripathi et al. 2024) and "rprop+" algorithm with a threshold value of 0.1. By using "e1071" package SVM model was developed in R software and was tuned using radial kernel function, grid search and cross-validation technique for robust evaluation and outlier detection of data (Mammone et al. 2009; Zhao et al. 2024). The gamma values ranged from 2^{-1} to 2^{-10}, while cost values were evaluated between 21 and 210. The best model was identified with a value 2 as cost parameter, 0.125 as gamma value, and 0.1 as epsilon value.

The GBM model was trained using the "caret" package (Demir et al. 2024). The developed model was tuned using "mlr" package and "tune ()" function to handle various types of data and the effects of outliers among predictor variables (Hastie et al. 2009; Kakati et al. 2022). The optimal model was created by configuring the parameters with 2,000 decision trees, an interaction depth of 1, a shrinkage parameter (Liu et al. 2024) of 0.01, minimum number of observations 5 and a bag fraction of 0.5 using threefold cross-validation to enhance robustness and prevent overfitting.

Validation of the trained models and predictions

The machine learning models were validated using data from test plots that was not involved in model building process. Among the four tested models, RF had higher prediction accuracy (RMSE: 0.27 t ha^{-1}, MAE: 0.22, MAPE: 5.32, NRMSE: 6.18%) compared to other three models i.e. GBM, ANN and SVM. The prediction accuracy of the four models was evaluated using RMSE, MAE, MAPE and NRMSE values, as shown in Table 5.

The predicted rice yield (t ha^{-1}) for the RF model in the test plots varied between 3.16 and 5.50, whereas the GBM, ANN, and SVM models ranged from 3.38 to 5.26, 3.36 to 5.50, and 3.19 to 5.88 (Fig. 3), respectively. Wang et al.

Table 5 Accuracy parameters of machine learning models used for prediction of rice yield using drone mounted RGB sensors

Model	RMSE ($t\ ha^{-1}$)	MAE	MAPE	NRMSE (%)
Support Vector Machine	0.69	0.57	14.34	15.81
Random Forest	0.27	0.22	5.32	6.18
Artificial Neural Network	0.65	0.56	13.60	14.89
Gradient Boosting Machine model	0.50	0.42	10.15	11.45

RMSE Root mean square error, *MAE* Mean absolute error, *MAPE* Mean absolute percentage error *NRMSE Normalised* root mean square error

(2022) used RGB sensor imagery to estimate rice aboveground biomass, with an R^2 of 0.81 and RMSE of 0.20. Similarly, Yamaguchi et al. (2020) used the combination of Deep learning methods along with RGB imageries and achieved enhanced prediction accuracy with R^2 of 0.96 and RMSE of 0.33 for Yield and LAI. Shiri et al. (2012) found that Random Forest outperformed other machine learning techniques in predicting corn yield, while Kuhnert et al. (2018) showed that Random Forest could accurately forecast wheat yield in Australia for complicated nonlinear problems, which reduces the influence of limited or anomalous data by randomly choosing samples to generate several classifiers (Pan et al. 2023).

Random Forest is not reliant on statistical assumptions and the need for extensive preprocessing, and it exhibits lower variance compared to a single decision tree (Zhu 2020). It effectively prevents overfitting and supports differential class weighting, outlier detection, and missing value imputation (Sharma et al. 2023). In contrast, some other machine learning algorithms face challenges such as insufficient sample sizes, data errors, and limitations in extrapolation, continuity, and smoothness when performing regression modeling on logged data (Ao et al. 2019). The accuracy of Random Forest can be attributed to its resilience against overfitting and its capacity to capture nonlinear relationships between environmental variables (e.g., temperature, rainfall) and crop yields (Kuradusenge et al. 2023). Similar to results of this study, other researchers (Singh and Singh 2017; Bhuyan et al. 2018; Sharma et al. 2023) also found RF model to be more accurate as compared to other ML methods (GBM, ANN and SVM) for estimating crop yield due to the robustness against overfitting (Zha et al. 2020). However, the efficacy of the ANN, GBM and SVM models in capturing complex relationships establishes them as a potent tool in crop yield predictive modelling, particularly when dealing with non-linear along with high-dimensional datasets due to their ability to capture higher-order interactions between predictors, as well as nonlinear relationships between predictors and target variable (Moghri et al. 2015; Jamshidi et al. 2016; Kelley et al. 2020; Sampaio et al. 2021; Shah et al. 2021; Alshari et al. 2023; Martinho 2023).

The differences in R-squared values across the models (Fig. 4) reflect their ability to capture and handle the complex, non-linear relationships that are typically encountered in predicting crop yield. The strong performance of RF model aligns with findings from previous studies (Jeong et al. 2016; Khaki & Wang 2019; Charoen and Mittrapiyanuruk 2018; Klompenburg et al. 2020), which also reported the effectiveness of RF in yield prediction. The studies have shown that RF models tend to outperform other machine learning techniques like Support Vector machine (SVM) and Polynomial Regression (PR) in predicting yields for crop such as maize and wheat (Bhatnagar and Gohain 2020; Zhang et al. 2020a, b; Li et al. 2023).

In this study, both the GBM and SVM ML models exhibited moderate performance. The ANN model exhibited the lowest performance, indicating its relatively low predictive accuracy. Other researchers (Chandrashekar and Sahin 2014; Hira and Gillies 2015; Sanz et al. 2018) have also utilized Random Forests and Support Vector Machines in the environmental and agronomical studies.

Similar studies that have explored models like SVM and ANN have also reported mixed results in terms of prediction (Matsumura et al. 2015; Crane-Droesch 2018; Ju et al. 2019; Ahmad et al. 2018). However, the superior performance of the RF model in this study, and in others, highlights its ability to handle complex, non-linear relationships and large datasets. The moderate predictions of the SVM and GBM models can be useful in specific contexts, but RF remains the most reliable model for robust rice yield prediction.

Variable importance

The variable importance of 17 vegetation indices (VIs) which were used in four models SVM, RF, ANN, and GBM, was analyzed based on the Gini impurity index (Godara et al. 2022) (Fig. 5). Among the top predictors of rice grain yield across these models were CI, I, and GMR. Notably, CI and I emerged as crucial in the RF and GBM models, while GMR played a key role in the ANN and GBM models. Moderately important indices included H, ExB, and ExR, contributing to prediction accuracy in models like SVM and ANN. On the other hand, indices such as g, VARI, and NDIg consistently ranked lower in importance, having minimal influence on yield prediction across all models.

The dominance of CI, I, and GMR as top predictors across multiple models is likely due to their close association with chlorophyll content, which is a key indicator of plant health and photosynthetic efficiency (Abdullah et al. 2019). Chlorophyll concentration reflects a plant's ability to carry out photosynthesis, making it an essential predictor of crop

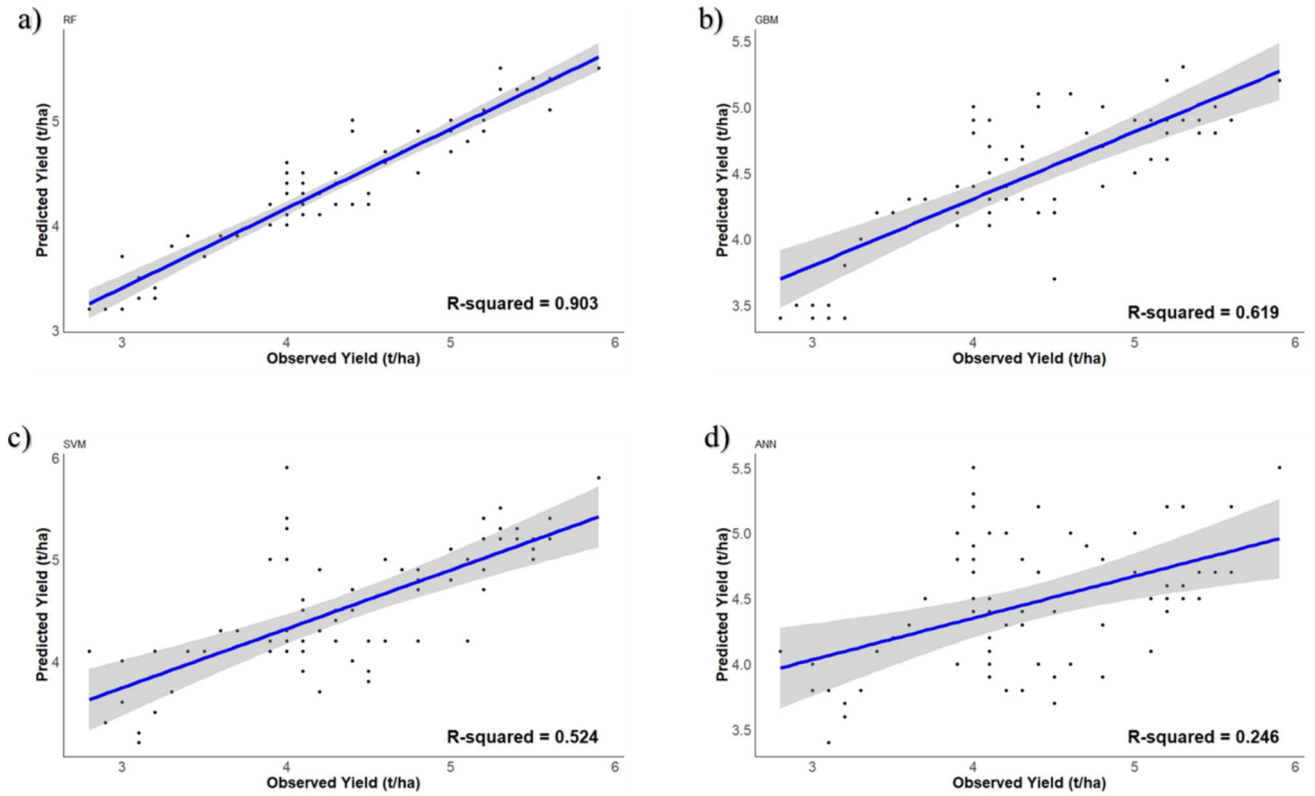


Fig. 4 Performance comparison of **a** RF (Random Forest), **b** GBM (Gradient Boosting Machine), **c** SVM (Support Vector Machine) **d** ANN (Artificial Neural Network) models for rice yield prediction based on RGB-derived vegetation indices

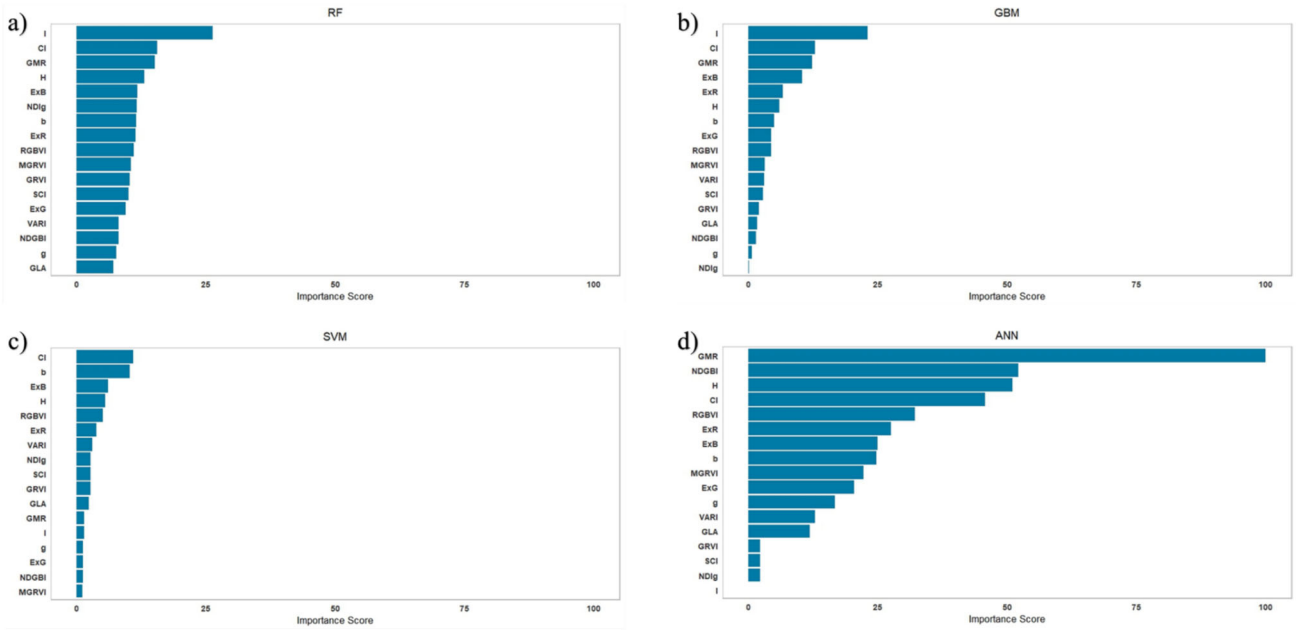


Fig. 5 Horizontal bar plot showing the contributions of predictor variables for **a** RF (Random Forest), **b** GBM (Gradient Boosting Machine), **c** SVM (Support Vector Machine), **d** ANN (Artificial Neural Network). *Note: ExR* Excess red index, *ExG* Normalized excess green index, *GLA* Green Leaf Algorithm, *GMR* Green minus red index, *SCI* Soil colour index, *MGRVI* Modified green red vegetation index, *I* Intensity, *GRVI* Green–Red Vegetation Index, *VARI* Visible Atmospherically Resistant Index, *NDGBI* Normalized Difference Green–Blue Index, *NDIg* Normalized red green difference index, *RGBVI* Red Green Blue Vegetation Index, *B* Normalized Blueness Intensity, *g* Green Percentage Index, *H* Hue, *CI* Coloration indices, *ExB* Excess blue vegetation indices, *r* Normalized red band, *G* Green band, *B* Blue band, *R* Red band

yield. As highlighted in previous studies, CI is highly sensitive to stress factors like drought, capturing early declines in chlorophyll and, consequently, crop vitality (Abdullah et al. 2019; Qiao et al. 2022). This aligns with our results, where CI ranked highly across models, emphasizing its value in yield prediction under varying environmental conditions. The intensity index (I) also emerged as a robust predictor, integrating reflectance across all visible bands (R, G, B) to provide a comprehensive measure of canopy health and structure. This aligns with findings by Ge et al. (2021a, b), which suggest that the intensity index is particularly effective in yield prediction due to its ability to capture spectral information from the entire canopy, offering insights into crop biomass and vigor.

Hue (H), which represents the dominant wavelength of reflected light, is highly sensitive to variations in plant health and chlorophyll content, making it a valuable predictor for yield, particularly in rice where subtle shifts in leaf color can indicate physiological stress or nutrient deficiency (Zakaluk 2008; Montanaro et al. 2023). Meanwhile, ExR and ExB, which capture early growth-stage data, help detect stress factors such as water or nutrient deficiency that can affect yield. Early-stage sensitivity, especially during tillering to booting stages, allows for the timely identification of growth conditions that influence final grain yield (Zhou et al. 2017; Killeen et al. 2024). Conversely, indices like VARI and NDIG exhibited lower importance, likely due to the saturation effect they experience during the late growth stages. As observed by Zhou et al. (2017), the increasing yellowing of leaves and the emergence of panicles at heading and filling stages reduced the predictive power of VARI. Similarly, NDIg reduced sensitivity during periods of dense canopy growth could explain its diminished importance, as it is less effective in capturing subtle stress signals compared to other indices (Qiao et al. 2022).

Conclusion

This study developed rice yield estimation models using drone-based RGB imagery, marking an innovative approach to yield prediction. By integrating vegetation indices extracted from RGB images, ML models: RF, SVM, ANN and GBM were developed for rice yield prediction. Among these, the Random Forest model showed the highest performance with an R^2 of 0.90 and an RMSE of 0.27 t ha^{-1}, demonstrating its robustness in accurately predicting rice yield from spectral inputs. These results underscore the practical utility of low-cost drone-mounted RGB sensors for real-time and non-invasive yield prediction, offering significant advantages for early decision-making in agriculture. Accurate pre-harvest yield forecasts can be a transformative tool for stakeholders, including insurance companies and policymakers, facilitating risk management and enabling more precise policy interventions based on actual crop conditions. While the study provides valuable insights, it remains a preliminary investigation conducted in tropical lowland rice using RGB sensors. Using low cost RGB sensors for detecting physiological changes across different rice crop growth stages still remains a challenge. Also, the generalizability of the developed ML models may be limited by region-specific crop phenology, different crop management practices and variable soil-climate interactions. More research work needs to be conducted incorporating multi-season and multi-location data to improve model accuracy, reliability, generalizability and reduce overfitting. Furthermore, the fusion of RGB, multispectral and thermal sensor data may be used to enhance the predictive accuracy. This study highlights the feasibility of using drone imagery to monitor crop growth and yield efficiently, providing an opportunity for its application in diverse agricultural systems. Drone-based technology has the potential to transform crop monitoring and yield forecasting, making it a powerful tool for improving agricultural productivity and management strategies.

Acknowledgements Authors acknowledge the logistical help provided by Director, ICAR-Central Rice Research Institute, Cuttack. This work was supported by the Indian Council of Agricultural Research, New Delhi and TIH iHub DRISHTI, IIT Jodhpur.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abdullah, H., Darvishzadeh, R., Skidmore, A.K., Heurich, M.: Sensitivity of Landsat-8 OLI and TIRS data to foliar properties of early-stage bark beetle (*Ips typographus*, L.) infestation. Remote Sens. **11**(4), 398 (2019). https://doi.org/10.3390/rs11040398

Ahmad, I., Saeed, U., Fahad, M., Ullah, A., Rahman, M.H.U., Judge, J.: Yield forecasting of spring maize using remote sensing and crop modeling in Faisalabad-Punjab. Pak. J. Indian Soc. Remote Sens. **46**, 1701–1711 (2018). https://doi.org/10.1007/s12524-018-0825-8

Akbarian, S., Xu, C.Y., Lim, S.: Analysis on the effect of spatial and spectral resolution of different remote sensing data in sugarcane crop yield study. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. **5**(3), 655–661 (2020).

Akinwande, M.O., Dikko, H.G., Samson, A.: Variance inflation factor: as a condition for the inclusion of suppressor variable(s) in regression analysis. Open J. Stat. **5**(7), 754 (2015)

Ali, A.M., Abouelghar, M., Belal, A.A., Saleh, N., Yones, M., Selim, A.I., Amin, M.E., Elwesemy, A., Kucher, D.E., Maginan, S., Savin, I.: Crop yield prediction using multi-sensors remote sensing. Egypt J. Remote Sens. Space Sci. **25**, 711–716 (2022). https://doi.org/10.1016/j.ejrs.2022.04.006

Alshari, E.A., Abdulkareem, M.B., Gawali, B.W.: Classification of land use/land cover using artificial intelligence (ANN-RF). Front Artif. Intell. (2023). https://doi.org/10.3389/frai.2022.964279

Amaratunga, V., Wickramasinghe, L., Perera, A., Jayasinghe, J., Rathnayake, U.: Artificial neural network to estimate the paddy yield prediction using climatic data. Math. Probl. Eng. **2020**, 8627824 (2020). https://doi.org/10.1155/2020/8627824

Ao, Y., Li, H., Zhu, L., Ali, S., Yang, Z.: The linear random forest algorithm and its advantages in machine learning-assisted logging regression modeling. J. Pet. Sci. Eng. **174**, 776–789 (2019). https://doi.org/10.1016/j.petrol.2018.11.067

Aranguren, M., Castellón, A., Aizpurua, A.: Crop sensor based non-destructive estimation of nitrogen nutritional status, yield, and grain protein content in wheat. Agriculture **10**(5), 148 (2020)

Bagheri, N., Ahmadi, H., Alavipanah, S.K.O.: Multispectral remote sensing for site-specific nitrogen fertilizer management. Braz. J. Agric. Res. **48**(10), 1394–1401 (2013). https://doi.org/10.1590/S0100-204X2013001000011

Bai, D., Li, D., Zhao, C., Wang, Z., Shao, M., Guo, B., Liu, Y., Wang, Q., Li, J., Guo, S., Wang, R.: Estimation of soybean yield parameters under lodging conditions using RGB information from unmanned aerial vehicles. Front. Plant Sci. **13**, 1012293 (2022)

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., Bareth, G.: Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. Int. J. Appl. Earth Obs. Geoinf. **39**, 79–87 (2015). https://doi.org/10.1016/j.jag.2015.02.012

Bhatnagar, R., Gohain, G.B.: Crop yield estimation using decision trees and random forest machine learning algorithms on data from Terra (EOS AM-1) & Aqua (EOS PM-1) satellite data. Mach. Learn. Data Min. Aerosp. Technol. 107–124 (2020).

Bhuyan, M., Bhuyan, M., Sarma, K.: Yield prediction of rice using machine learning algorithms. Int. J. Comput. Sci. Eng. **6**(3), 51–55 (2018)

Bishop, C.M.: Pattern Recognition and Machine Learning. Springer (2006)

Blackmer, T.M., Schepers, J.S., Varvel, G.E., Meyer, G.E.: Analysis of aerial photography for nitrogen stress within corn fields. Agron. J. **88**, 729–733 (1996)

Breiman, L.: Bagging predictors. Mach. Learn. **24**(2), 123–140 (1996)

Breiman, L.: Random forest. Mach. Learn. **45**, 5–32 (2001)

Cen, H., Wan, L., Zhu, J., Li, Y., Li, X., Zhu, Y., Weng, H., Wu, W., Yin, W., Xu, C.: Dynamic monitoring of biomass of rice under different nitrogen treatments using a lightweight UAV with dual image-frame snapshot cameras. Plant Methods **15**, 32 (2019)

Chandrashekar, G., Sahin, F.: A survey on feature selection methods. Comput. Electr. Eng. **40**(1), 16–28 (2014)

Chang, X.: Rice yield prediction based on deep learning. Environ. Monit. Assess. **195**, 490–496 (2024)

Charoen-Ung, P., Mittrapiyanuruk, P.: Sugarcane yield grade prediction using random forest and gradient boosting tree techniques. In: 2018 15th Int. Jt. Conf. Comput. Sci. Softw. Eng. (JCSSE), 1–6 (2018).

Chen, Z., Wang, F., Zhang, P., Ke, C., Zhu, Y., Cao, W., Jiang, H.: Skewed distribution of leaf color RGB model and application of skewed parameters in leaf color description model. Plant Methods **26**, 16–23 (2020)

Chu, Z., Yu, J.: An end-to-end model for rice yield prediction using deep learning fusion. Comput. Electron. Agric. **174**, 105471 (2020)

Crane-Droesch, A.: Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. Environ. Res. Lett. **13**(11), 114003 (2018)

Dahikar, S.S., Rode, S.V.: Agricultural crop yield prediction using artificial neural network approach. Int. J. Innov. Res. Electr. Electron. Instrum. Control Eng. **2**(1), 683–686 (2014)

Das, B., Sahoo, R.N., Pargal, S., Krishna, G., Gupta, V.K., Verma, R., Viswanathan, C.: Measuring leaf area index from colour digital image of wheat crop. J. Agrometeorol. **18**(1), 22–28 (2016)

Das, S.R., Nayak, B.K., Dey, S., Sarkar, S., Chatterjee, D., Saha, S., Sarkar, D., Pradhan, A., Saha, S., Nayak, A.K.: Potential soil organic carbon sequestration vis-a-vis methane emission in lowland rice agroecosystem. Environ. Monit. Assess. **195**, 1099 (2023)

Demir, S., Dedeoğlu, M., Başayiğit, L.: Yield prediction models of organic oil rose farming with agricultural unmanned aerial vehicles (UAVs) images and machine learning algorithms. Remote Sens. Appl. Soc. Environ. **33**, 101131 (2024)

Dev, V.A., Eden, M.R.: Formation lithology classification using scalable gradient boosted decision trees. Comput. Chem. Eng. **128**, 392–404 (2019)

Escadafal, R., Belghit, A., Ben-Moussa, A.: Indices spectraux pour la télédétection de la dégradation des milieux naturels en Tunisie aride. In: Guyot, G. (ed) Actes du 6ème Symposium international sur les mesures physiques et signatures en télédétection, Val d'Isère (France), 17–24 Janvier, 253–259 (1994).

Feng, H., Tao, H., Li, Z., Yang, G., Zhao, C.: Comparison of UAV RGB imagery and hyperspectral remote-sensing data for monitoring winter wheat growth. Remote Sens. **14**(15), 3811 (2022)

Ferchichi, A., Abbes, A.B., Barra, V., Farah, I.R.: Forecasting vegetation indices from spatio-temporal remotely sensed data using deep learning-based approaches: a systematic literature review. Ecol. Inform. **68**, 101552 (2022)

Fieuzal, R., Sicre, C.M., Baup, F.: Estimation of corn yield using multi-temporal optical and radar satellite data and artificial neural networks. Int. J. Appl. Earth Obs. Geoinf. **57**, 14–23 (2017)

Fu, H., Wang, C., Cui, G., She, W., Zhao, L.: Ramie yield estimation based on UAV RGB images. Sensors **21**(2), 669 (2021)

Gandhi, N., Armstrong, L.J., Petkar, O., Tripathy, A.K.: Rice crop yield prediction in India using support vector machines. In: Proc. 13th Int. Jt. Conf. Comput. Sci. Softw. Eng. (JCSSE), Khon Kaen, Thailand, 1–5 (2016a).

Gandhi, N., Petkar, O., Armstrong, L.J.: Rice crop yield prediction using artificial neural networks. In: 2016 IEEE Technol. Innov. ICT Agric. Rural Dev. (TIAR), 105–110 (2016b).

García-Martínez, H., Flores-Magdaleno, H., Ascencio-Hernández, R., Khalil-Gardezi, A., Tijerina-Chávez, L., Mancilla-Villa, O.R., Vázquez-Peña, M.A.: Corn grain yield estimation from vegetation indices, canopy cover, plant density, and a neural network using multispectral and RGB images acquired with unmanned aerial vehicles. Agriculture **10**(7), 277 (2020)

Ge, H., Ma, F., Li, Z., Du, C.: Grain yield estimation in rice breeding using phenological data and vegetation indices derived from UAV images. Agronomy **11**(12), 2439 (2021a)

Ge, H., Xiang, H., Ma, F., Li, Z., Qiu, Z., Tan, Z., Du, C.: Estimating plant nitrogen concentration of rice through fusing vegetation indices and color moments derived from UAV-RGB images. Remote Sens. **13**, 1620 (2021b)

Geipel, J., Link, J., Claupein, W.: Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. Remote Sens. **6**, 10335–10355 (2014)

Gerard, B., Buerkert, A.: Estimation of spatial variability in pearl millet growth with non-destructive methods. Exp. Agric. **37**, 373–389 (2001)

Gitelson, A.A., Dall'Olmo, G., Moses, W., Rundquist, D.C., Barrow, T., Fisher, T.R., Gurlin, D., Holz, J.: A simple semi-analytical model for remote estimation of chlorophyll-an in turbid waters: validation. Remote Sens. Environ. **112**(9), 3582–3593 (2008)

Godara, P., Verma, S., Kumari, S., Kumar, S.: Importance of variable using Gini index and discriminant score in Indian mustard genotypes. J. Agric. Res. Technol. **47**, 100–105 (2022)

Gong, Y., Duan, B., Fang, S., Zhu, R., Wu, X., Ma, Y., Peng, Y.: Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. Plant Methods **14**, 70 (2018)

Guo, Y., Wang, H., Wu, Z., Wang, S., Sun, H., Senthilnath, J., Wang, J., Bryant, C.R., Fu, Y.: Modified red-blue vegetation index for chlorophyll estimation and yield prediction of maize from visible images captured by UAV. Sensors **20**(18), 5055 (2020)

Gupta, S.D., Ibaraki, Y., Pattanayak, A.K.: Development of a digital image analysis method for real-time estimation of chlorophyll content in micropropagated potato plants. Plant Biotechnol. Rep. **7**, 91–97 (2013)

Han, R., Wen, Y., Bai, L., Liu, J., Choi, J.: Rate splitting on mobile edge computing for UAV-aided IoT systems. IEEE Trans. Cogn. Commun. Netw. **6**(4), 1193–1203 (2020)

Hasan, U., Sawut, M., Chen, S.: Estimating the leaf area index of winter wheat based on unmanned aerial vehicle RGB-image parameters. Sustainability **11**, 6829 (2019)

Hastie, T., Tibshirani, R., Friedman, J.: Boosting and additive trees. In: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, pp. 337–387 (2009).

He, Y., Dong, J., Liao, X., Sun, L., Wang, Z., You, N., et al.: Examining rice distribution and cropping intensity in a mixed single- and double-cropping region in south China using all available Sentinel 1/2 images. Int. J. Appl. Earth Obs. Geoinf. **101**, 102351 (2021)

Hertel, T.W., Burke, M.B., Lobell, D.B.: The poverty implications of climate-induced crop yield changes by 2030. Glob. Environ. Change **20**, 577–585 (2010)

Herzig, P., Borrmann, P., Knauer, U., Klück, H., Kilias, D., Seiffert, U.: Evaluation of RGB and multispectral unmanned aerial vehicle (UAV) imagery for high-throughput phenotyping and yield prediction in barley breeding. Remote Sens. **13**, 2670 (2021)

Hira, Z.M., Gillies, D.F.: A review of feature selection and feature extraction methods applied on microarray data. Adv. Bioinf. **2015**, 198363 (2015)

Hu, H., Liu, H., Zhang, H., Zhu, J., Yao, X., Zhang, X., Zheng, K.: Assessment of chlorophyll content based on image color analysis, comparison with SPAD-502. In: Proceedings of the 2nd International Conference on Information Engineering and Computer Science, Wuhan, China, 25–26 December 2010.

Huber, F., Inderka, A., Steinhage, V.: Leveraging remote sensing data for yield prediction with deep transfer learning. Sensors **24**(3), 770 (2024)

Hunt, E.R., Doraiswamy, P.C., McMurtrey, J.E., Daughtry, C.S.T., Perry, E.M., Akhmedov, B.: A visible band index for remote sensing leaf chlorophyll content at the canopy scale. Int. J. Appl. Earth Obs. Geoinf. **21**, 103–112 (2013)

Jamshidi, S., Yadollahi, A., Ahmadi, H., Arab, M.M., Eftekhari, M.: Predicting in vitro culture medium macro-nutrients composition for pear rootstocks using regression analysis and neural network models. Front. Plant Sci. **7**, 198363 (2016)

Jeong, J.H., Resop, J.P., Mueller, N.D., Fleisher, D.H., Yun, K., Butler, E.E., Timlin, D.J., Shim, K.-M., Gerber, J.S., Reddy, V.R., et al.: Random forests for global and regional crop yield predictions. PLoS ONE **11**, 1–15 (2016)

Jhajharia, K., Mathur, P.: Prediction of crop yield using satellite vegetation indices combined with machine learning approaches. Adv. Space Res. **72**(9), 3998–4007 (2023)

Jia, L., Chen, X., Zhang, F., Buerkert, A., Roemheld, V.: Optimum nitrogen fertilization of winter wheat based on color digital camera images. Commun. Soil Sci. Plant Anal. **38**, 1385–1394 (2007)

Ju, S., Lim, H., Heo, J.: Machine learning approaches for crop yield prediction with MODIS and weather data. In: Proceedings of the 40th Asian Conference on Remote Sensing: Progress of Remote Sensing Technology for Smart Future, ACRS, Daejeon, Republic of Korea, pp. 1–4 (2019).

Kakati, N., Deka, R.L., Das, P., Goswami, J., Khanikar, P.G., Saikia, H.: Forecasting yield of rapeseed and mustard using multiple linear regression and ANN techniques in the Brahmaputra Valley of Assam, Northeast India. Theor. Appl. Climatol. **150**(3), 1201–1215 (2022)

Kawashima, S., Nakatani, M.: An algorithm for estimating chlorophyll content in leaves using a video camera. Ann. Bot. **81**, 49–54 (1998)

Kazantsev, N., Islam, N., Zwiegelaar, J., Brown, A., Maull, R.: Data sharing for business model innovation in platform ecosystems: from private data to public good. Technol. Forecast. Soc. Chang. **192**, 122515 (2023)

Kefauver, S.C., Vicente, R., Vergara-Díaz, O., Fernandez-Gallego, J.A., Kerfal, S., Lopez, A., Melichar, J.P., Serret Molins, M.D., Araus, J.L.: Comparative UAV and field phenotyping to assess yield and nitrogen use efficiency in hybrid and conventional barley. Front. Plant Sci. **8**, 1733 (2017)

Kelley, J., McCauley, D., Alexander, G.A., Gray, W.F., Siegfried, R., Oldroyd, H.J.: Using machine learning to integrate on-farm sensors and agro-meteorology networks into site-specific decision support. Trans. ASABE **63**(5), 1427–1439 (2020)

Khaki, S., Wang, L.: Crop yield prediction using deep neural networks. Front. Plant Sci. **10**, 621 (2019)

Killeen, P., Kiringa, I., Yeap, T., Branco, P.: Corn grain yield prediction using UAV-based high spatiotemporal resolution imagery, machine learning, and spatial cross-validation. Remote Sens. **16**(4), 683 (2024)

Kim, J.H.: Multicollinearity and misleading statistical results. Korean J. Anesthesiol. **72**(6), 558 (2019)

Kuhn, M., Johnson, K.: Applied predictive modeling, vol. 26, p. 13. Springer, New York (2013)

Kuhnert, M., Yeluripati, J., Smith, P., Hoffmann, H., van Oijen, M., Constantin, J., Jansson, P.E.: Impact of model structure on the accuracy of wheat yield forecasts in a data-scarce region. Eur. J. Agron. **101**, 1–12 (2018)

Kumar, C., Mubvumba, P., Huang, Y., Dhillon, J., Reddy, K.: Multi-stage corn yield prediction using high-resolution UAV multispectral data and machine learning models. Agronomy **13**(5), 1277 (2023). https://doi.org/10.3390/agronomy13051277

Kuradusenge, M., Hitimana, E., Hanyurwimfura, D., Rukundo, P., Mtonga, K., Mukasine, A., Uwitonze, C., Ngabonziza, J., Uwamahoro, A.: Crop yield prediction using machine learning models: case of Irish potato and maize. Agriculture **13**(1), 225 (2023)

Kuwata, K., Shibasaki, R.: Estimating corn yield in the United States with MODIS EVI and machine learning methods. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. **3**, 131–136 (2016)

Lee, K.J., Lee, B.W.: Estimation of rice growth and nitrogen nutrition status using color digital camera image analysis. Eur. J. Agron. **48**, 57–65 (2013)

Li, B., Xu, X., Zhang, L., Han, J., Bian, C., Li, G.: Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. ISPRS J. Photogramm. Remote Sens. **162**, 161–172 (2020)

Li, R., Wang, D., Zhu, B., Liu, T., Sun, C., Zhang, Z.: Estimation of grain yield in wheat using source–sink datasets derived from RGB and thermal infrared imaging. Food Energy Secur. **12**(2), e434 (2023)

Li, S., Liu, X., Tian, Y., Zhu, Y., Cao, Q.: Comparison RGB digital camera with active canopy sensor based on UAV for rice nitrogen status monitoring. In: Proceedings of the 7th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Hangzhou, China, 6–9 August (2018)

Li, S., Yuan, F., Ata-UI-Karim, S.T., Zheng, H., Cheng, T., Liu, X., Tian, Y., Zhu, Y., Cao, W., Cao, Q.: Combining color indices and

textures of UAV-based digital imagery for rice LAI estimation. Remote Sens. **11**, 1763 (2019). https://doi.org/10.3390/rs11151763

Li, W., Niu, Z., Chen, H., Li, D., Wu, M., Zhao, W.: Remote sensing estimation of winter wheat biomass and yield based on multitemporal vegetation indices and crop height using airborne RGB and multispectral sensors. PLoS ONE **9**(11), e112791 (2014)

Liu, K., Li, Y., Han, T., Yu, X., Ye, H., Hu, H., Hu, Z.: Evaluation of grain yield based on digital images of rice canopy. Plant Methods **15**, 28 (2019)

Liu, S., Wang, J., Wang, H., Wu, Y.: Post-processing of hydrological model simulations using the convolutional neural network and support vector regression. Hydrol. Res. **53**(4), 605–621 (2022a)

Liu, Y., Chen, Y., Wen, M., Lu, Y., Ma, F.: Accuracy comparison of estimation on cotton leaf and plant nitrogen content based on UAV digital image under different nutrition treatments. Agronomy **13**, 1686 (2023)

Liu, Y., Hatou, K., Aihara, T., Kurose, S., Akiyama, T., Kohno, Y., Lu, S., Omasa, K.: A robust vegetation index based on different UAV RGB images to estimate SPAD values of naked barley leaves. Remote Sens. **13**, 686 (2021)

Liu, Y., Wang, S., Wang, X., Chen, B., Chen, J., Wang, J., Huang, M., Wang, Z., Ma, L., Wang, P., Amir, M.: Exploring the superiority of solar-induced chlorophyll fluorescence data in predicting wheat yield using machine learning and deep learning methods. Comput. Electron. Agric. **192**, 106612 (2022b)

Liu, Z., Ju, H., Ma, Q., Sun, C., Lv, Y., Liu, K., Wu, T., Cheng, M.: Rice yield estimation using multi-temporal remote sensing data and machine learning: a case study of Jiangsu, China. Agriculture **14**(4), 638 (2024). https://doi.org/10.3390/agriculture14040638

Louhaichi, M., Borman, M.M., Johnson, D.E.: Spatially located platform and aerial photography for documentation of grazing impacts on wheat. Geocarto Int. **16**(1), 65–70 (2001)

Lussem, U., Bolten, A., Gnyp, M.L., Jasper, J., Bareth, G.: Evaluation of RGB-based vegetation indices from UAV imagery to estimate forage yield in grassland. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. **XLII–3**, 1215–1219 (2018)

Maimaitijiang, M., Sagan, V., Sidike, P., Maimaitiyiming, M., Hartling, S., Peterson, K.T., Maw, M.J.W., Shakoor, N., Mockler, T., Fritschi, F.B.: Vegetation Index Weighted Canopy Volume Model (CVMVI) for soybean biomass estimation from Unmanned Aerial System-based RGB imagery. ISPRS J. Photogramm. Remote Sens. **151**, 27–41 (2019)

Mammone, A., Turchi, M., Cristianini, N.: Support vector machines. Wiley Interdiscip. Rev. Comput. Stat. **1**(3), 283–289 (2009)

Mariotto, I., Thenkabail, P.S., Huete, A., Slonecker, E.T., Platonov, A.: Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HyspIRI mission. Remote Sens. Environ. **139**, 291–305 (2013)

Martinho, V.J.P.: Fertiliser cost prediction in European Union farms: Machine-learning approaches through artificial neural networks. Open Agric. **8**(1) (2023)

Matsumura, K., Gaitan, C.F., Sugimoto, K., Cannon, A.J., Hsieh, W.W.: Maize yield forecasting by linear regression and artificial neural networks in Jilin. China. J. Agric. Sci. **153**, 399–410 (2015)

Meena, S.V., Dhaka, V.S., Sinwar, D.: Exploring the role of Vegetation indices in Plant diseases Identification. In: 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), pp. 372–377. IEEE (2020, November).

Meyer, G.E., Hindman, T.W., & Laksmi, K.: Machine vision detection parameters for plant species identification. In: Precision Agriculture and Biological Quality (Boston, MA: International Society for Optics and Photonics), pp. 327–336 (1999).

Moghri, M., Shamaee, H., Shahrajabian, H., Ghannadzadeh, A.: The effect of different parameters on mechanical properties of PA-6/clay nanocomposite through genetic algorithm and response surface methods. Int. Nano Lett. **5**(3), 133–140 (2015)

Mohite, J.D., Sawant, S.A., Pandit, A., Agrawal, R., Pappula, S.: Soybean crop yield prediction by integration of remote sensing and weather observations. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. **48**, 197–202 (2023)

Montanaro, G., Petrozza, A., Rustioni, L., Cellini, F., Nuzzo, V.: Phenotyping key fruit quality traits in olive using RGB images and back propagation neural networks. Plant Phenomics **5**, 0061 (2023)

Murthy, C.S., Poddar, M.K., Choudhary, K.K., Pandey, V., Srikanth, P., Ramasubramanian, S., Senthil, K.G.: Paddy crop insurance using satellite-based composite index of crop performance. Geomat. Nat. Hazards Risk **13**, 310–336 (2022)

Nahiyoon, S.A., Ren, Z., Wei, P., Li, X., Li, X., Xu, J., Yan, X., Yuan, H.: Recent development trends in plant protection UAVs: a journey from conventional practices to cutting-edge technologies—a comprehensive review. Drones **8**, 457 (2024)

Naik, S., GH, H.N., Rao, S.G.: Prediction of wheat yield by using UAV RGB drone imagery and advanced machine learning techniques. Comput. Electron. Agric. **183**, 106114 (2023)

Nakasagga, S., Adak, A., Murray, S.C., Rooney, W.L., Hoffmann, Jr. L., Wilde, S., Lindsey, R., Nabukalu, P., Cox, S.: Prediction of regrowth and biomass of perennial sorghum using unoccupied aerial systems. Crop Sci. **62**(6) (2022)

Niu, Q., Feng, H., Yang, G., Li, C., Yang, H., Xu, B., Zhao, Y.: Monitoring plant height and leaf area index of maize breeding material based on UAV digital images. Trans. Chin. Soc. Agric. Eng. **34**(5), 73–82 (2018)

Noureldin, N.A., Aboelghar, M.A., Saudy, H.S., Ali, A.M.: Rice yield forecasting models using satellite imagery in Egypt. Egypt. J. Remote Sens. Space Sci. **16**(1), 125–131 (2013)

Özreçberoğlu, N., Kahramanoğlu, İ: Mathematical models for the estimation of leaf chlorophyll content based on RGB colours of contact imaging with smartphones: a pomegranate example. Folia Hortic. **32**, 57–67 (2020)

Pan, W., Wang, X., Sun, Y., Wang, J., Li, Y., Li, S.: Karst vegetation coverage detection using UAV multispectral vegetation indices and machine learning algorithm. Plant Methods **19**(1), 7 (2023)

Pazhanivelan, S., Kumaraperumal, R., Shanmugapriya, P., Sudarmanian, N.S., Sivamurugan, A.P., Satheesh, S.: Quantification of biophysical parameters and economic yield in cotton and rice using drone technology. Agriculture **13**(9), 1668 (2023)

Peng, Y., Li, Y., Dai, C., Fang, S., Gong, Y., Wu, X., Zhu, R., Liu, K.: Remote prediction of yield based on LAI estimation in oilseed rape under different planting methods and nitrogen fertilizer applications. Agric. for. Meteorol. **271**, 116–125 (2019)

Prasad, N.R., Patel, N.R., Danodia, A.: Crop yield prediction in cotton for regional level using random forest approach. Spat. Inf. Res. **29**, 195–206 (2021)

Probst, P., Boulesteix, A.L., Bischl, B.: Tunability: importance of hyperparameters of machine learning algorithms. J. Mach. Learn. Res. **20**(53), 1–32 (2019)

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S.: A modified soil adjusted vegetation index. Remote Sens. Environ. **48**(2), 119–126 (1994). https://doi.org/10.1016/0034-4257(94)90134-1

Qiao, L., Zhao, R., Tang, W., An, L., Sun, H., Li, M., Wang, N., Liu, Y., Liu, G.: Estimating maize LAI by exploring deep features of vegetation index map from UAV multispectral images. Field Crops Res. **289**, 108739 (2022)

Qiu, Z., Ma, F., Li, Z., Xu, X., Du, C.: Development of prediction models for estimating key rice growth variables using visible and NIR images from unmanned aerial systems. Remote Sens. **14**(6), 1384 (2022)

R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, https://www.R-project.org/ (2023)

Richardson, A.D., Jenkins, J.P., Braswell, B.H., Hollinger, D.Y., Ollinger, S.V., Smith, M.L.: Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. Oecologia **152**(2), 323–334 (2007)

Ripley, B.: Feed-forward neural networks and multinomial log-linear models. nnet package, version 7.3–6 (2013). http://www.stats.ox.ac.uk/pub/MASS4

Ruwanpathirana, P.P., Sakai, K., Jayasinghe, G.Y., Nakandakari, T., Yuge, K., Wijekoon, W.M.C.J., Priyankara, A.C.P., Samaraweera, M.D.S., Madushanka, P.L.A.: Evaluation of sugarcane crop growth monitoring using vegetation indices derived from RGB-based UAV images and machine learning models. Agronomy **14**(9), 2059 (2024)

Saberioon, M.M., Amin, M.S.M., Anuar, A.R., Gholizadeh, A., Wayayok, A., Khairunniza-Bejo, S.: Assessment of rice leaf chlorophyll content using visible bands at different growth stages at both the leaf and canopy scale. Int. J. Appl. Earth Obs. Geoinf. **32**, 35–45 (2014)

Sampaio, P.S., Almeida, A.S., Brites, C.M.: Use of artificial neural network model for rice quality prediction based on grain physical parameters. Foods **10**(12), 3016 (2021)

Sanches, G.M., Duft, D.G., Kölln, O.T., Luciano, ACd.S., De Castro, S.G.Q., Okuno, F.M., Franco, H.C.J.: The potential for RGB images obtained using unmanned aerial vehicles to assess and predict yield in sugarcane fields. Int. J. Remote Sens. **39**(15–16), 5402–5414 (2018)

Sánchez-Sastre, F., da Veiga, N., Alte, M.S., Ruiz-Potosme, M., Carrión-Prieto, P., Marcos-Robles, L., Navas-Gracia, L., Martín-Ramos, P.: Assessment of RGB vegetation indices to estimate chlorophyll content in sugar beet leaves in the final cultivation stage. AgriEngineering, **2**(1) (2020)

Sanz, H., Valim, C., Vegas, E., Oller, J.M., Reverter, F.: SVM-RFE: Selection and visualization of the most relevant features through non-linear kernels. BMC Bioinform. **19**(1), 432 (2018)

Shah, M.I., Alaloul, W.S., Alqahtani, A., Aldrees, A., Musarat, M.A., Javed, M.F.: Predictive modeling approach for surface water quality: development and comparison of machine learning models. Sustainability **13**(14), 7515 (2021)

Shahhosseini, M., Hu, G., Archontoulis, S.V.: Forecasting corn yield with machine learning ensembles. Front. Plant Sci. **11**, 1120 (2020)

Sharma, S.K., Sharma, D.P., Gaur, K.: Crop yield predictions and recommendations using random forest regression in 3A agroclimatic zone, Rajasthan. J. Data Acquis. Process. **38**(2), 1635 (2023)

Shigeto, K., Nakatani, M.: An algorithm for estimating chlorophyll content in leaves using a video camera. Ann. Bot. **81**(1), 49–54 (1998)

Shiri, J., Keshavarzi, A., Kisi, O., Iturraran-Viveros, U., Mousavi Rizi, S.M.: Modeling daily reference evapotranspiration using available and estimated solar radiation data by adaptive neuro-fuzzy and neural network computing methods. J. Hydrol. **414**, 417–429 (2012)

Singh, S., Singh, R.: Predicting crop yield using machine learning techniques: a case study of wheat, paddy and mustard crops in Punjab. Indian J. Sci. Technol. **10**(11), 1–7 (2017)

Sitienei, M., Anapapa, A., Otieno, A.: Random Forest Regression in Maize Yield Prediction. Asian J. Prob. Stat. **23**(4), 43–52 (2023)

Song, C., Sang, J., Zhang, L., Liu, H., Wu, D., Yuan, W., Huang, C.: Adaptiveness of RGB-image derived algorithms in the measurement of fractional vegetation coverage. BMC Bioinform. **23**(1), 358 (2022)

Steyerberg, E.W., Vergouwe, Y.: Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur. Heart J. **35**(29), 1925–1931 (2014)

Sumesh, K.C., Ninsawat, S., Som-Ard, J.: Integration of RGB-based vegetation index, crop surface model and object-based image analysis approach for sugarcane yield estimation using unmanned aerial vehicle. Comput. Electron. Agric. **180**, 105903 (2021)

Tanaka, Y., Watanabe, T., Katsura, K., Tsujimoto, Y., Takai, T., Tanaka, T.S.T., Kawamura, K., Saito, H., Homma, K., Mairoua, S.G., Ahouanton, K.: Deep learning enables instant and versatile estimation of rice yield using ground-based RGB images. Plant Phenomics **5**, 0073 (2023)

Tian, Z., Zhang, Y., Liu, K., Li, Z., Li, M., Zhang, H., Wu, J.: UAV remote sensing prediction method of winter wheat yield based on the fused features of crop and soil. Remote Sens. **14**, 5054 (2022)

Tilly, N., Aasen, H., Bareth, G.: Fusion of plant height and vegetation indices for the estimation of barley biomass. Remote Sens. **7**, 11449–11480 (2015)

Tripathi, R., Tripathy, B.R., Jena, S.S., Swain, C.K., Mohanty, S., Sahoo, R.N., Nayak, A.K.: Prediction of rice yield using sensors mounted on unmanned aerial vehicle. Agric. Res. 1–11 (2024).

Tucker, C.J.: Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. **8**(2), 127–150 (1979)

Valdivieso-Ros, C., Alonso-Sarria, F., Gomariz-Castillo, F.: Effect of the synergetic use of sentinel-1, sentinel-2, lidar and derived data in land cover classification of a semiarid Mediterranean area using machine learning algorithms. Remote Sens. **15**(2), 312 (2023)

Van Klompenburg, T., Kassahun, A., Catal, C.: Crop yield prediction using machine learning: a systematic literature review. Comput. Electron. Agric. **177**, 105709 (2020)

Verma, V., Vishal, B., Kohli, A., Kumar, P.P.P.: Systems-based rice improvement approaches for sustainable food and nutritional security. Plant Cell Rep. **40**, 1–16 (2021)

Volpato, L., Pinto, F., González-Pérez, L., Thompson, I.G., Borém, A., Reynolds, M., Gérard, B., Molero, G., Rodrigues, F.A., Jr.: High throughput field phenotyping for plant height using UAV-based RGB imagery in wheat breeding lines: feasibility and validation. Front. Plant Sci. **12**, 591587 (2021)

Wan, L., Cen, H., Zhu, J., Li, Y., Zhu, Y., Sun, D., Weng, H., He, Y.: Combining UAV-based vegetation indices, canopy height and canopy coverage to improve rice yield prediction under different nitrogen levels. In: 2019 ASABE Annual International Meeting. Am Soc Agric Biol Eng, **1** (2019). https://doi.org/10.13031/aim.201900564

Wang, F., Yi, Q., Hu, J., Xie, L., Yao, X., Xu, T.: Combining spectral and textural information in UAV hyperspectral images to estimate rice grain yield. Int. J. Appl. Earth Obs. Geoinf. **102**, 102397 (2021)

Wang, Y., Wang, D., Zhang, G., Wang, J.: Estimating nitrogen status of rice using the image segmentation of G-R thresholding method. Field Crops Res. **149**, 33–39 (2013)

Wang, Z., Ma, Y., Chen, P., Yang, Y., Fu, H., Yang, F., Ma, J.: Estimation of rice aboveground biomass by combining canopy spectral reflectance and unmanned aerial vehicle-based red green blue imagery data. Front. Plant Sci. **13**, 903643 (2022)

Wei, Q., Li, L., Ren, T., Wang, Z., Wang, S., Li, X., Cong, R., Lu, J.: Diagnosing nitrogen nutrition status of winter rapeseed via digital image processing technique. Sci. Agric. Sin. (2015).

Widjaja Putra, T.B., Soni, P.: Enhanced broadband greenness in assessing chlorophyll a and b, carotenoid, and nitrogen in Robusta coffee plantations using a digital camera. Precis. Agric. **19**, 238–256 (2018)

Wilding, L.P.: Spatial variability: its documentation, accommodation, and implication to soil surveys. In: Nielsen, D.R., Bouma, J. (eds.) Soil Spatial Variability, Pudoc, Wageningen, Netherlands (1985).

Wilke, N., Siegmann, B., Postma, J.A., Muller, O., Krieger, V., Pude, R., Rascher, U.: Assessment of plant density for barley and wheat using UAV multispectral imagery for high-throughput field phenotyping. Comput. Electron. Agric. **189**, 106380 (2021)

Woebbecke, D.M., Meyer, G.E., Von Bargen, K., Mortensen, D.A.: Color indices for weed identification under various soil, residue, and lighting conditions. Trans. ASAE **38**(1), 259–269 (1995)

Wu, W.: Intelligent Monitoring and System Development of Diagnostic Index of the Wheat Seedling Growth. *YangZhou University* (2016).

Xue, J., Su, B.: Significant remote sensing vegetation indices: a review of developments and applications. J. Sens. **2017**(1), 1353691 (2017)

Yamaguchi, T., Tanaka, Y., Imachi, Y., Yamashita, M., Katsura, K.: Feasibility of combining deep learning and RGB images obtained by unmanned aerial vehicle for leaf area index estimation in rice. Remote Sens. **13**(1), 84 (2020)

Yang, L., Shami, A.: On hyperparameter optimization of machine learning algorithms: theory and practice. Neurocomputing **415**, 295–316 (2020)

Yang, S., Li, L., Fei, S., Yang, M., Tao, Z., Meng, Y., Xiao, Y.: Wheat yield prediction using machine learning method based on UAV remote sensing data. Drones **8**, 284 (2024)

Yue, J., Yang, G., Tian, Q., Feng, H., Xu, K., Zhou, C.: Estimate of winter-wheat above-ground biomass based on UAV ultrahigh-ground-resolution image textures and vegetation indices. ISPRS J. Photogramm. Remote Sens. **150**, 226–244 (2019)

Zakaluk, R.F.: Ground-based RGB imaging to determine the leaf water potential of potato plants (2008).

Zakaluk, R., Ranjan, R.S. Predicting the leaf water potential of potato plants using RGB reflectance. Can. Biosyst. Eng. J. 50, 7.1–7.12 (2008)

Zeng, L., Peng, G., Meng, R., Man, J., Li, W., Xu, B., Lv, Z., Sun, R.: Wheat yield prediction based on unmanned aerial vehicles-collected red–green–blue imagery. Remote Sens. **13**(15), 2937 (2021)

Zha, H., Miao, Y., Wang, T., Li, Y., Zhang, J., Sun, W., Feng, Z., Kusnierek, K.: Improving unmanned aerial vehicle remote sensing-based rice nitrogen nutrition index prediction with machine learning. Remote Sens. **12**(2), 215 (2020)

Zhang, Z., Flores, P., Igathinathane, C., Naik, D.L., Kiran, R., Ransom, J.K.: Wheat lodging detection from UAS imagery using machine learning algorithms. Remote Sens. **12**, 1838 (2020a)

Zhang, M., Zhou, J., Sudduth, K.A., Kitchen, N.R.: Estimation of maize yield and effects of variable-rate nitrogen application using UAV-based RGB imagery. Biosyst. Eng. **189**, 24–35 (2020b)

Zhao, Y., Zhang, W., Liu, X.: Grid search with a weighted error function: hyper-parameter optimization for financial time series forecasting. Appl. Soft Comput. **154**, 111362 (2024)

Zhang, Y., Jiang, Y., Xu, B., Yang, G., Feng, H., Yang, X., Yang, H., Liu, C., Cheng, Z., Feng, Z.: Study on the estimation of leaf area index in rice based on UAV RGB and multispectral data. Remote Sens. **16**(16), 3049 (2024)

Zhang, Y., Li, M., Xu, Z., Zhang, X., Wang, M.: Investigation of crop nitrogen content based on image processing technologies. Proc. SPIE **5909**, 539–547 (2005)

Zhou, H., Yang, J., Lou, W., Sheng, L., Li, D., Hu, H.: Improving grain yield prediction through fusion of multi-temporal spectral features and agronomic trait parameters derived from UAV imagery. Front. Plant Sci. **14**, 1217448 (2023)

Zhou, X., Zheng, H.B., Xu, X.Q., He, J.Y., Ge, X.K., Yao, X., Cheng, T., Zhu, Y., Cao, W.X., Tian, Y.C.: Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. ISPRS J. Photogramm. Remote Sens. **130**, 246–255 (2017)

Zhu, T.: Analysis on the applicability of the random forest. J. Phys. Conf. Ser. **1607**(1), 012123 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.